2

USAFSAM-TR-90-33

REVIEW OF AUDITION LITERATURE: SELECTION OF ACOUSTIC SIGNALS FOR USE IN THE SYNTHESIS OF AUDITORY SPACE

Don C. Teas, Ph.D.

KRUG Life Sciences
San Antonio Division
405 West Nakoma
San Antonio, TX 78216

December 1990

Final Report for Period August 1989 - August 1990

Approved for public release; distribution is unlimited.

Prepared for
USAF SCHOOL OF AEROSPACE MEDICINE
Human Systems Division (AFSC)
Brooks Air Force Base, TX 78235-5301

NOTICES

This final report was submitted by KRUG Life Sciences, San Antonio Division, 405 West Nakoma, San Antonio, Texas, under contract F33615-89-C-0603-02, job order 7930-20-03 with the USAF School of Aerospace Medicine, Human Systems Division, AFSC, Brooks Air Force Base, Texas. Dr. Kent K. Gillingham (USAFSAM/VNB) was the Laboratory Project Scientist-in-Charge.

This report has been reviewed and is approved for publication.

Kent K. Gillingham

KENT K. GILLINGHAM, M.D., Ph.D.
Project Scientist

William F. Storm

WILLIAM F. STORM, Ph.D.
Supervisor

George Schwender MD

GEORGE E. SCHWENDER, Colonel, USAF, MC, CFS
Commander

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION Unclassified	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY 2b DECLASSIFICATION/DOWNGRADING SCHEDULE	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release, distribution is unlimited.
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S) USAFSAM-TR-90-33
6a NAME OF PERFORMING ORGANIZATION KRUG Life Sciences San Antonio Division 6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION USAF School of Aerospace Medicine (VNB)
6c ADDRESS (City, State, and ZIP Code) 405 West Nakoma San Antonio, TX 78216	7b ADDRESS (City, State, and ZIP Code) Human Systems Division (AFSC) Brooks AFB TX 78235-5301
8a NAME OF FUNDING/SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F33615-89-C-0603-02
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO 62202F; PROJECT NO 7930; TASK NO 20; WORK UNIT ACCESSION NO 03

11 TITLE (Include Security Classification)
Review of Audition Literature. Selection of Acoustic Signals for Use in the Synthesis of Auditory Space

12 PERSONAL AUTHOR(S)
Teas, Don C.

13a. TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Final	FROM 89/08 TO 90/08	1990, December	36

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Audition, Sound Localization, Auditory Objects, Auditory Research, Auditory Localization
20	01		
06	04		

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

The purpose of this review is to aid pilot orientation. It provides information about topics of current interest in psychoacoustics that seem particularly relevant to the problem of synthesizing auditory space and presenting information within it Included is a historical background with an orientation toward contemporary psychoacoustic information. Appendix A lists, with descriptors, the journal articles reviewed. Appendix B is a bibliography.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Kent K. Gillingham, M.D., Ph.D.	22b TELEPHONE (Include Area Code) (512)536-3521 22c OFFICE SYMBOL USAFSAM/VNB

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

CONTENTS

REVIEW OF AUDITION LITERATURE: SELECTION OF ACOUSTIC SIGNALS FOR USE IN THE SYNTHESIS OF AUDITORY SPACE

BACKGROUND

Man's perceptions became objects of proper study only in the 19th century. The early investigators were "natural philosophers," trained broadly in physics, philosophy and, often, in medicine. There were many difficult questions that often arose from philosophy; the empirical answers were sometimes surprising. For example, the "armchair" conclusion that thought must occur at the speed of light was made absurd in 1850 when Helmholtz measured the speed of the nerve impulse at 27 meters/sec (4). Even if the nerve impulse itself was not "thought," the concept was that thought is somehow related to nerve impulses. Since a nerve impulse travels at a finite speed, its product must also be finite. From 1879 forward, the development of laboratories for psychological research provided an academic home for the empirical study of sensation and perception.

The perception of one's position in space, up and down, with respect to other objects and places, and also the perception of the position of one's head, arms, and legs captured the interest of these early workers. Human sensory capabilities and motor skills are easily observed: we run to targets, avoid objects, throw and catch balls in the air, all with great accuracy as if we maintained a detailed three-dimensional map by which our muscle controller makes decisions about its output signals. We are also aware of our actions. The rich interaction between philosophy and the new experimental psychology in the 19th century led to the formulation of broad principles and questions about the relation between sensory input and subsequent behavior. Even though the tools for empirical study were limited, the contributions of the "new" experimental psychology to the problem of spatial relations have endured.

We understand today that the senses are "input ports" which provide data to central locations that process neural information. We know that the receptor systems and central processing sites are in anatomical registration so that in the auditory system, for example, the distribution of frequency along the cochlea is replicated in central nuclei. The auditory system is, therefore, said to be <u>tonotopically</u> organized. Even before the anatomy describing the projection of sensory receptor systems to central nuclei was available, Lotze, in 1852, reasoned that there were <u>Local Signs</u>, i.e., signatures, to represent a code for every spot on the skin. The same basic notion holds for the visual and the postural receptor systems: locations in space are projected to locations on the retina, and each angular or linear direction in space is represented by a semicircular canal or otolith organ, respectively. Thus, the physical world is first mapped onto receptor surfaces and the spatial relations on those surfaces are retained in their central projections. In this way the physical world is represented in the neuroanatomy of the sensory system.

The proprioceptors, including the vestibular receptor systems, provide information about the positions of head, arms, legs, feet and hands. Because of this representation we can act, i.e., move about, on the basis of sensory information from the environment. The relative state and position of the muscles and joints is delivered to the nerve fibers that carry the information back to (other) central processors to be integrated with the most recent sensory information about the outside world. Updated information about the outside world is then delivered to the effectors for the next moment of action. There must be a continuous integration of sensory information about the outside world with information about current body position in order to output the next command for effector placement. At each instant, we would expect that the sensory inflow (outside world) must be evaluated for its match to a desired value (stored template), which, in turn, must be derived from an objective; e.g., throw the ball to a target. The target must be designated prior to the first of a series of movements, then completed when the ball strikes its target, since eye movement must follow the ball after it has left the hand to confirm that the muscle action produced the expected result. The small task of throwing a ball includes many of the questions that the early experimental-physiological psychologists tried to address.

Beginnings of Empirical Study

Hearing presented a problem to the generalists studying spatial awareness in the 19th century because, unlike the retina or the skin, or the specific assignments for each semicircular canal, the receptor for sound has room to represent only frequency and (perhaps) intensity, and none for outside space, yet listeners can readily localize sound sources. The spatial attribute of sound was difficult to assign to one auditory receptor. Other attributes of sound, i.e., pitch, loudness, and timbre were studied by the investigators using only cumbersome resonators, monochords and tuning forks. In a first-order sense, the physical correlate for pitch was known to be frequency, for loudness, the intensity of sound and, for timbre, variation in the number of tone sources. That sound required a medium for conduction, that the velocity of sound was about 1130 ft/sec, and that pitch varied with frequency were all known as well as the relations between length, tension and mass for a stretched string. The presence of overtones, divided into the fundamental and harmonics, was also recognized. Two insights set the scientific stage for Helmholtz's resonance theory of hearing which dominated research for many decades. In 1822, Fourier, studying heat, found that any continuous function could be analyzed into a series of sine waves that varied in period, amplitude and phase. Thus, the stretched string which vibrated in parts (harmonics) as well as over its entire length (fundamental) was a physical system of which Fourier's Theorem was an analog. The analysis of the system could be made by using resonators of different frequencies to identify the frequencies corresponding to the vibrations of the string in parts. The fundamental corresponds to the displacement of the entire string, the second harmonic (first partial) corresponds to the vibration of the string in two halves, etc. In 1843, Ohm argued that the ear can distinguish the frequencies produced by the vibrations of the stretched string in its parts, thus announcing Ohm's Law of Hearing. Ohm's analytic principle has been amply supported in studies of the identification of distortion products in the ear. In 1863, Helmholtz published his theory of hearing, Sensations of Tone. He incorporated the anatomical knowledge that had accumulated since the invention of the microscope. The organ of Corti was known to contain hair cells, supporting cells, and was located on the basilar membrane. The fundamental mechanism of resonance of stretched strings seemed to fit the structures along the basilar membrane. Pitch was determined by the place along the membrane at which displacement occurred, loudness by the amplitude of displacement. Somehow the auditory nerve fibers at the resonance peak, where the displacement was greatest, were stimulated. Those nerve impulses from that location was the code for perceiving pitch. The principle of resonance provided the frequency analysis needed to incorporate Ohm's law.

Helmholtz's theory focused on the attributes of sound that are supported by a single ear, i.e., monaural rather than binaural. He inadvertently slowed the acceptance of interaural phase difference as a cue for localization of sinusoids, however. He was unable to determine any effect of phase changes on pitch, loudness or timbre in his stimuli and, therefore, quite logically, ignored phase in his theory. The consequence was that his authority -- even though it should not have -- led others to deny that phase effects were detectable, even when interaural phase differences were demonstrated to be useful for localization. Only in recent years has the effect of phase changes been recognized as producing changes in timbre, despite Helmholtz's observations (28).

Later, Bekesy (2) pointed out that the cochlear partition, including the basilar membrane, the organ of Corti and the tectorial membrane were all displaced with acoustic stimulation and that the basilar membane was not under tension as Helmholtz's resonance theory required. He further showed that the cochlear partition represented a system that exhibited traveling waves moving in one direction regardless of the location at which stimulation occurred. The broad amplitude maximum of the traveling wave is located near the base of the cochlea for high frequencies, and moves to the apex as frequency decreases. Increases in stimulus intensity produce increases in the amplitudes of displacement along the cochlear partition; consequently, there is also some modification of place of stimulation that could excite nerve fibers.

Early Studies of Sound Localization

The recognition that the disparity of stimulation at pairs of receptor systems (e.g., the two retinas and the two cochleas) provides the cues for visual depth and auditory localization, did not come for vision until 1775 and for hearing, except for casual "armchair" mention, until 1846 (5). Wheatstone invented the stereoscope in 1833, thus isolating the retinal disparity cue and synthesizing visual depth. The analogue for hearing was not to appear for over a hundred years until stereophonic sound in the 1940s, and even then the synthesis of auditory disparity was not as singular an experience as is produced by synthesizing retinal disparity. Only within the last decade has sufficient computer power been generally available to synthesize or to reconstitute the complex auditory stimuli that produce the rich perceptions that direct experience generates. Indeed, only within the present century was the vacuum tube developed and were researchers able to control the frequency and amplitude of oscillatory signals with precision.

As described by Boring (5) the first report of sound localization was by E.H. Weber in 1846. He noted that two watches, placed on each side of an observer, could both be heard at once and their location recognized. We know that the two ticks or tocks must have been heard separately since continuous, similar sounds from two different azimuth locations fuse into an apparent single source with its location dependent upon the relative intensities of the two sounds but usually at a position between the two real sources. In 1877 Lord Rayleigh reported observations on sound localization carried out on his lawn. In the center of a circle of his assistants, he localized their different voices to within a few degrees. Tuning forks were localized with less success. Rayleigh knew, of course, that the shorter the wavelength, the greater the sound shadow produced by the head from a lateral position of the source. Since he had trouble localizing tones from low frequency tuning forks, Rayleigh concluded that interaural intensity differences provided the cue for localization. He also pointed out that the same interaural difference can exist in the rear plane and thus, front-back reversals are likely, but there is no confusion among azimuth angles in the frontal plane.

In the same year Sylvanus Thompson, who became Lord Rutherford, observed 'binaural beats', a phenomenon heard when one low tone is led to one ear and another, slightly mistuned, is led to the other ear. One hears a waxing and waning of intensity of an auditory image that moves within the head. If the frequencies are further separated, the beating diminishes and one hears two different sounds at the ears. Thompson reported later that the position of a sound heard through tubes to the ears changed when the phase of one tuning fork was altered. Later, in 1907, Rayleigh proposed a phase theory after duplicating Thompson's earlier observations. Once Rayleigh had proposed phase as a cue--in opposition to the Helmholtz legacy--other workers then described studies that had been suppressed, due to the Helmholtz denial of phase "perception".

The term, phase, is appropriately used for a continuous sinusoid. The ticks or tocks of E.H. Weber's watches were discontinuous with abrupt onsets, similar to clicks. When impulsive sounds arrive at the ears at the same time, the source is heard in the median plane, dead ahead. As the interval between the times of arrival of the sound at the two ears increases, the source is heard to move from the median plane toward the leading ear. We now know that the time interval for just detecting a difference from center is 10 μs for an optimal sound in an optimal environment (18). Von Hornbostel and Wertheimer reported in 1920 that 30 μs were required.

Pre-Contemporary Experiments

During the development of vacuum tube technology most of the knowledge about hearing was captured in Helmholtz's resonance theory. The "Theory of Hearing" was interpreted to be theory about cochlear function; the central representation of the attributes of sound was not addressed except to refer to "the sensorium". Pitch depended on the place of stimulation, depending in turn upon the tension and mass of the cochlear strands. Multiple frequencies could exist along the basilar membrane since locations would resonate according to the frequencies contained in the stimulus. Observers could hear these components, thus substantiating the analytical nature of the receptor system in the manner suggested by the Fourier Theorem. Localization was not a salient feature of the theory since it required registration of stimulus differences at the two ears and Helmholtz's concern was to account for those attributes which were present for monaural stimulation, principally pitch, but with a bow toward loudness. Even though it was outside the Helmholtz definition of auditory theory, localization also benefited from the increased stimulus control available with vacuum tube technology.

Simultaneous Masking

As auditory research absorbed the new technology, new demonstrations and tests of ideas and deductions from the resonance theory were made. The most significant was the observation by H. Fletcher in 1940 that the intensity of a band of random noise at which a sinusoidal signal was masked was equal to the intensity of the sinusoid. That is, if the noise contains the same energy as the signal, and the two stimuli are present simultaneously, the signal is replaced in one's perception and only the noise is audible. Let the signal be a sinusoid of 1000 Hz. We begin with a noise containing sinusoidal components from 200 - 5000 Hz and present it through earphones at a comfortable listening level. We adjust the intensity of the sinusoid, the signal, so that it can be detected only half the time. Now, the limits of the noise are reduced from 200 to, say, 400 Hz and from 5000 to 3000 Hz for a bandwidth of 2600 Hz. The signal remains at the same intensity and retains its detectability. Progressive narrowing of the noise band leaves the signal detectability about the same until the width is near 200 Hz, say, from 900 - 1100 Hz. Further narrowing of the noise band produces an increase in the signal's detectability; signal intensity must be reduced to restore masking. The bandwidth at which the signal detectability increases by a criterion amount is taken as the Critical Band (CB), and is interpreted as a "functional unit length" along the cochlear partition. One can also plot the CB inversely; i.e., beginning with a sinusoid, add frequencies and the loudness of the sound will remain the same until frequencies outside the critical band are added, at which point loudness increases.

Fletcher's observations provided the auditory community with a psychometric tool to study hearing, using the observer as a meter, a null instrument. The perception of a signal could be measured in terms of its replacement by noise, i.e., one perceptual quality could be substituted for another, quantitatively. Only the noise in a narrow bandwidth around the signal, the CB, is effective as the masker and its center frequency follows the frequency of the signal. One assumes that the CB is passed by a filter surrounding the signal frequency. As frequency is increased, the width of the critical band filter increases. The shape of the CB filter has been studied extensively (29). The masking experiment has remained a psychoacoustic tool and the CB has become a reference point. The related concept of a filter has more generality and has been used to describe physiological as well as psychological responses. As yet, there has been little direct study of physiological events related to complex acoustic signals.

The CB was orginally defined by monaural, simultaneous presentation of masker and signal. The masking noise was continuous and the signal was pulsed. The signal can be turned on and off gradually to minimize onset and offset transients which could spread energy across several critical bands, thus obscuring the interpretation of the masked threshold. The rigorous operational definition of the masked threshold led to close agreement of masked thresholds among laboratories. The CB seems to be a rock-solid construct describing an important parameter of hearing. Hearing includes parameters that extend beyond the conditions defining the monaural CB, but the masking paradigm has proved sufficiently general to accommodate a wide range of experimental questions.

Temporal Masking

In particular, a class of experiments called temporal masking has isolated the effects produced when the masker precedes the signal, and when the masker follows the signal, forward and backward masking, respectively. The interpretation of forward masking is that a segment of the auditory system holds, for a time, the effect of the masker; i.e., there has been insufficient time from masker offset for recovery of that segment of the auditory system, and the response to the signal is reduced. Thus, forward masking is studied as a function of the time between the offset of the masker and the onset of the signal. Recovery occurs in about 100 msec. Backward masking is more difficult to interpret while retaining the usual order of causality. Presumably the masker, usually stronger than the signal, elicits neural activity with a shorter latency than the signal, thus the excitation due to the masker "catches up" with the weaker excitation from the signal. The time over which backward masking occurs is about 50 ms. The parameters of temporal masking have relevance for any sequential auditory stimulation such as speech.

Ordinarily one would expect that the auditory filter might be measured most effectively by masking with tones. However, other phenomena such as beats and distortion products can interfere with the detection of a tonal signal. One can minimize the occurrence of beats by using short tones, but at the expense of broadening the spectrum. With forward masking the problem of interaction between two tonal signals is avoided, and the effect of the masker can be measured by determining the masked threshold for a probe tone. In such an experiment, the masker frequency is varied, and the intensity at each frequency is adjusted to mask the probe tone which is set to a sensation level (SL) within 10 to 20 dB of threshold. One finds that the intensity of the masker required to mask the low-level probe is least when its frequency is near the probe signal. As frequency deviates from the probe, more intensity is required. In this way a curve that resembles the pass band of a filter is determined. The curve derived with forward masking is narrower than that found with simultaneous masking.

Binaural Masking

The principle of masking was extended to binaural discrimination. The parameters of masker and signal become more complicated for binaural stimulation. In particular, the signal and the masker can have different phase relations with respect to the two ears. The noise can be "in-phase", i.e., each tympanic membrane moving inward or outward at the same time, or in "phase opposition", i.e., one tympanic membrane moving outward while the other is moving inward. Similarly, the signal can be in interaural phase agreement or phase opposition. The noise and signal are independently variable. The experimenter still measures the masked threshold, but now there are more stimulus conditions than in the monaural case. For the binaural condition in which the noise and the signal are both in phase agreement, the masked threshold is the same as for the monaural case. Hirsh (17) showed that the masked threshold obtained in the binaural condition for which the signal (S) is in phase opposition (180 deg) while the noise (N) is in phase agreement (0 degrees) was about 11 dB lower than the monaural or binaural phase-agreement condition. Subsequent work has shown that for low frequencies the NoSπ condition (interaural phase relations: noise at 0° and signal at 180°) produces a binaural Masking Level Difference (MLD) of 15 dB. Other combinations of binaural noise and masker conditions produce smaller MLDs.

From the largest MLD of 15 dB at 250 Hz, there is a decrease to the vanishing point at about 4000 Hz. The role of interaural phase in determining masked threshold and the low frequencies at which phase is effective suggests that the underlying physiological mechanism for the MLD may also serve for sound localization. The site at which the MLD is generated must be central, i.e., where the inputs from the two ears interact. Thus, the binaural CB, wider than the monaural, may reflect neural processing at a central rather than peripheral site.

Localization of Sound

The roles of interaural time differences and interaural intensity differences as cues for sound localization, shown to be important by the early work of Rayleigh and of Thompson in 1877 (5), survived in the study by Stevens and Newman (33) for which there was adequate stimulus control. Stevens and Newman (33) showed that low frequencies, up to about 1000 Hz, were accurately localized and frequencies above about 3000 Hz were also accurately localized. Between these two limits there was a frequency region for which localization was poor. They suggested that at low frequencies the interaural phase differences provided an accurate cue while for the high frequencies which produced a sound shadow, interaural intensity differences provided the cue. These observations referred only to sound sources in the horizontal plane, i.e., azimuth angle.

Although interaural time and intensity are important to both binaural masking and localization, questions about the two cues cannot, it seems, be exactly overlaid. In binaural masking experiments, the stimuli form a sound image; the signal can appear in the "intracranial" perceptual space in a different place from the noise. Highly-trained observers can detect two images, one related to interaural intensity difference, the other to interaural time or phase difference (24,25). When the two cues are put into opposition, observers can report on each image (12). However, either cue will move the sound image produced by a click from the center of the head; an image offset by a small interaural intensity difference can be returned to the middle of the head with a small time difference favoring the opposite ear (38). For larger interaural differences between simple stimuli, two images can be discerned. Such separate analyses of interaural time and intensity differences can be done only by delivering stimuli via earphones. The differences between stimulating the binaural system via earphones and via external sound sources is recognized by the terms, lateralization, for earphones, and localization, for spatially-located sound sources (30).

In the early 1970s the attributes of hearing were thought to depend upon much the same stimulus parameters as was the case for the decade of the 1930s, even though great increments of detail about discrimination among sounds had been added. The knowledge base about the parameters of the auditory system, both psychoacoustic and physiological, had vastly increased. The method of study implied that the effects measured by sinusoids might be summed to predict the effects produced by more complicated signals such as speech and other complex sounds. Binaural masking was an intriguing window into the auditory system that might be related to phenomena such as the selection of one signal out of many -- the cocktail party effect -- wherein a listener can pick out of babble one particular voice for attention. Even so, interaural time and intensity differences were thought to be the basis for binaural phenomena, whether localization, lateralization or binaural masking.

If the early 1970s was a consolidation period, during which the status quo was strengthened, the late 1970s and the 1980s was a time for questioning that steady state of auditory theory. In the description below we review recent psychoacoustic work with complex acoustic signals, much of which does not require binaural stimulation. We will then describe the binaural work with complex signals.

CONTEMPORARY RESEARCH

Two reports of experiments by Watson and his colleagues (34, 35) have provided an important basis for contemporary developments in the study of discrimination among acoustic signals. In their first report they showed that detection of changes in intensity or frequency of sine components in a tonal sequence varied with position in the sequence. In their second report the investigators showed that such discriminations depended directly on stimulus uncertainty. Watson and his colleagues suggested that at minimal stimulus uncertainty, one could study the physiological resolving power of the auditory system, as, for example, represented by CB experiments, and, as stimulus uncertainty increased, one could also study how humans process conditional acoustic inputs. For example, speech, like any other sound, must first be processed acoustically, but the listener may then have a series of choices, with uncertainty among them reduced by context.

The incorporation of stimulus uncertainty into contemporary psychoacoustics has proceeded quickly. In his recent book, Profile Analysis (11), Green describes how his studies of the effect of stimulus uncertainty upon the detection of intensity increments (beginning about 1980) led to the study of the spectral shape of complex signals and the experimental isolation of unexpected capabilities of auditory discrimination. An important instrumental advantage for the experimental control of stimulus uncertainty has been the use of computer-generated stimuli. The computer can select rapidly among stored sinusoids and combine them to produce complex signals that vary in component frequencies and intensities and output them through high-speed digital to analog converters to earphones for subject's decisions. Rules for choosing component frequencies and their intensities can be constructed to guide the subjects' decision rules. The classic Profile Analysis experiment will be used below to introduce some of the principal findings that are emerging from the contemporary study of discrimination among complex acoustic signals.

Profile Analysis

The profile which is analyzed by the subject in this category of experiments is the pattern of the components of the complex signal, i.e., the relative energies among the components. On a horizontal axis representing frequency, each component has a location; on the vertical axis, each component has a height, representing its energy. Thus, there is a vertical line for each component frequency that reaches some height on the vertical axis. When the vertical lines all end at the same ordinate value, the profile is flat. The stimulus is produced by combining all the component frequencies into a single voltage waveform delivered to the subject's earphones. The subject hears a complex signal, perhaps, 100 ms in duration, with a flat spectrum. A second spectrum is now prepared, differing from the first by an increment in the middle frequency component. The middle component terminates at a higher ordinate value than the other components. Within 250 ms or so from the end of the first signal, the second complex signal is output. The second signal resembles the first, but the increment in the intensity of the middle component may change the sound. When detecting a difference, the subject indicates whether the signal occurred in the first or second interval. The noise+signal (the second profile) may occur in either interval. The single component of the noise+signal stimulus is incremented until the subject chooses, with some predetermined probability, that stimulus as the one containing the signal. The amount of that increment is the detection threshold of the subject for the alteration in the stimulus profile.

If asked to describe the difference between the two complex sounds, noise (flat profile) and the noise+signal, it is unlikely that the subject could identify the increment in the middle component of the complex stimulus. Instead, the subject detects a difference in quality between the two complex sounds. Since the stimuli are constructed from basic components, the effect upon discrimination of variation in each feature of the complex signals can be studied. The number of components can be varied, different components can be selected, the component carrying the increment can be varied, etc. However, if the number of components is reduced to one, the essence of the Profile Analysis experiment is lost. In this case, the detection of the intensity increment is a successive comparison between the single component in each interval. With two components or more in the stimulus profile, the subject is said to make simultaneous comparisons of intensities among the component frequencies. The number of components has been varied from 1 to 20. As the number of components is increased, the detection thresholds require larger increments in the signal component. Green (11) lists the following variations:

i. For the case illustrated above, the signal and the masker were both fixed and the uncertainty was minimal.

ii. The signal can remain fixed and the frequency components of the masker can be varied from trial to trial.

iii. The increment is added to any component of the set of components; thus, the signal frequency varies, but the masker components remain fixed.

iv. The increment is added to any frequency component, randomly as in iii, and the masker frequencies are also changed from trial to trial, as for ii. Thus, the signal and masker are both random with respect to frequency, and uncertainty is relatively high.

Of the four conditions, the subjects require most intensity in the increment for the conditions described in iv; i.e., both signal and masker randomized, and required the least increment for i, with neither signal nor noise randomized, i.e.,the least uncertainty. Subjects performed more poorly for ii, the randomized masker, than for iii, the randomized signal. As Green points out, this last finding is surprising since the masker should have little to do with the energy in the CB surrounding the incremented middle component. Another randomization was made in the intensity of the profile (the height of all components) from interval to interval within a range. Variation over a range of as much as 30 dB or so increased the detection threshold by no more than 2 dB. The "roving" intensity of the profile removed all but the relative differences among its components and forced the subjects to base their detection on that feature of the stimuli. Since the threshold was perturbed only by a small amount, the conclusion is that the auditory system can discriminate signals on the basis of relative differences among components.

Among the many observations in the context of Profile Analysis, one of the most interesting is the effect of the number of components surrounding the signal component, perhaps because it stands in some contrast to the concept of the CB. For the case with only three components, a middle one, the signal, and two adjacent ones, detection of an increment in the signal was found to improve as a function of the frequency range spanned by the two side components. With the signal at the middle component, the effect of the number of components was studied, for a maximum of 21. The increment required to detect the signal was at a minimum for 11 components, spaced at equal log intervals. If additional components had entered the CB surrounding the signal, one would expect that the signal would be more difficult to detect. However, conventional masking has absolutely nothing to contribute to the interpretation of an increase in detectability, i.e., a lowered threshold, with the addition of masker energy remote from the CB. Indeed, when the CB is invaded by additional energy from crowded components as their numbers increase, conventional masking does occur and the detectability of the signal decreases. The finding that detectability improves when energy outside the CB is present is consistent with the inference that subjects assay spectral shape by making simultaneous comparisons among frequency components. Although the explanation for 11 component frequencies being an optimum number is not clear, other studies using different paradigms have also shown that off-signal frequencies improve detection of signals. In particular, steady-state noises shaped to resemble vowels can be discriminated from babble-noise (9).

Comodulation Masking Release (CMR)

The extra-CB effects seen in Profile Analysis have been studied with other experimental strategies. Hall, Haggard and Fernandes (15) showed that the threshold for a signal in a noise band could be decreased if the noise band surrounding the signal and another band with a different but nearby center-frequency were modulated identically. The comodulation of the two noise bands is usually accomplished by multiplying a narrow band of low frequency noise, for example, 0-50 Hz, by a sinusoid to translate the center-frequency to mid-range, then filtering the unwanted bands to leave the "flanking band" and the band surrounding the signal. One interpretation of the improvement in detectability of the sinusoidal signal is that the vector addition of signal and noise produces an event different in the masking band from that in the flanking band. Thus the difference in temporal variation in the two envelopes (signal+noise band vs. flanking band) is detected. McFadden (23) showed that detection was not locked specifically to the comodulation of the two noise bands by creating experimental conditions in which detection was improved for random rather than comodulated noise. Instead of a sinusoid as signal, McFadden (23) used a narrow band of noise. There were as many as four narrow-band noises flanking the signal band. Detection was improved for the condition in which the signal band was not correlated with the flanking bands, a reversal of the expected CMR result. The phenomenological explanation is, of course, that the contrast of the signal band with the background is important for detection, and, in McFadden's study, contrast was greatest between the noise and signal for the uncorrelated case. A contrast interpretation may also account for the small or absent CMR effects for signal frequencies below 1000 Hz. Richards (31) found that subjects could discriminate between correlated bands of noise when the center frequencies were less than an octave apart,and when their separation was greater than 1000 Hz. For noise bands with center frequencies separated by an octave or at frequencies as low as 350 Hz, subjects could not discriminate between noise bands. McFadden's (23) result suggests the interpretation that, for those two cases, the perceptual contrast between bands to be discriminated was minimal. The octave is twice the frequency and would be expected to duplicate some of the temporal variation. The rates of variation in acoustic pressure for noise bands with low center-frequencies overlap with pressure variations due to the random amplitude fluctuations of a narrow band of noise. In both cases the contrast due to the experimental manipulation is reduced.

The amount of threshold reduction produced by comodulation masking release has varied among studies since its first demonstration. The initial study by Hall, Haggard & Fernandes (15) showed a threshold decrease of 10 dB. McFadden (22) studied the amount of CMR 1) as the intensity of the "flanking" or cue band was varied, 2) as signal duration was varied, 3) for differences in times of onset of the masker and cue bands, and 4) in a forward masking paradigm. The CMR was largest, about 10 dB, when the masker and cue bands were equal, at 70 dB SPL. The CMR averaged about 7 dB for increases in signal duration from 75 to 375 ms. A CMR maximum of 8 dB was observed at 0.8 ms difference between the onsets of the cue and masker bands for 75-Hz bandwidths, while for 100-Hz bandwidths, the maximum was 6 dB. Finally, McFadden (22) reported a "residual" CMR of 3 dB under forward masking conditions which he later (23) accounted for from considerations other than CMR.

There is probably agreement among the groups that have studied CMR that there is, indeed, an "across-frequency" effect on detection of a signal in noise by a flanking, comodulated band. There is disagreement concerning how large an effect can be attributed to such a mechanism. Schooneveldt and Moore (32) would attribute 10-15 dB of the total CMR to within-CB phenomena, i.e., phase effects, and but 2-4 dB to across-frequency listening. Hall and Grose (14) suggest that CMR is "multiply-cued". Whether the CMR magnitude also depends on differences in the way the complex signals for these experiments are generated is unknown.

Cohen and her co-workers (7,19) and Hall, Cokely and Grose (13) studied the possibility that the monaural release from masking due to comodulation of masker and cue bands is related to the binaural release from masking caused by interaural phase disparities between the signal and masker. Hall et al. (13) found that four of their six subjects were able to combine the interaural phase cue and the comodulation cue to achieve greater release from masking than either cue provided separately. However, the data from two of their subjects did not show that capability. Cohen and Schubert (7) reported a binaural CMR smaller than the expected binaural masking level difference. The comparisons among stimulus conditions and the alterations in detectability that are expected from these combinations are not clear. However, the interaural phase effects exist at frequencies below about 1000 Hz and the comodulation effects depend upon narrow bandwidths which produce envelope variations at low frequencies. Perhaps the release from masking produced by both of these procedures depends upon low frequencies. The stimulus manipulations at low frequencies alter the detectability of the signal in noise, perhaps also modifying the salience of the signal.

Modulation

In Profile Analysis, Comodulation Masking Release, and also in Binaural Release from Masking, it is the threshold of detection which is measured, i.e., the change in intensity required to detect the signal at some predetermined probability. Because it is the intensity increment from some suprathreshold loudness that is to be detected in Profile Analysis, we can determine that the subject perceives a change in quality, or timbre, of the sound rather than an increase in loudness of a single component frequency. Some investigators have studied suprathreshold signals directly in an attempt to determine the stimulus correlates for the perceptual segregation of complex acoustic signals. In the description of Profile Analysis the pattern of frequencies for the stimuli could be described in spectral terms; viz., for signal, all the components along the frequency axis reached the same value on the ordinate. The components were all combined into one voltage waveform and presented to the subject. Suppose that, for some group of components, the height along the ordinate is varied during the time of presentation, i.e., amplitude modulated (AM). The components receiving AM will stand in perceptual relief from those not being modulated. Other stimulus modifications will also produce perceptual segregation of components, e.g., differences in location, in loudness, in moment of onset, in duration, in pitch, etc. Simultaneous changes in many of these parameters probably contribute to the perceptual separation of one voice from many.

Yost and his coworkers (37, 39) have studied the effect of variations in the parameters of Sinusoidal Amplitude Modulation (SAM) upon the segregation of complex sounds into auditory groups or "objects". Detection of SAM is best for modulation rates below 50 Hz in that the depth of modulation required is least. Modulation depth must be increased about 4 times from that required at 20 Hz in order to detect the presence of modulation at the rate of 200 Hz. At the low rates of SAM, where detection is best, the segregation of two auditory carriers by amplitude modulation is most fragile, that is, there must be relatively large differences in modulation depth between the two carriers in order to perceive them as separate. Detection of a change in modulation rate requires an increase of 10%.

McAdams (21) has reported on the segregation effects of frequency modulation (FM) using synthesized vowels. Each vowel was presented simultaneously for three different fundamental frequencies. The separations among vowel formant frequencies were maintained for the shifts in pitch. The fundamental frequencies of the vowels, /a/, /i/, or /o/, either target or background, were frequency modulated. Corresponding to the degree of perceived certainty that a designated vowel was present in the three-vowel complex, the subject moved a slider along a scale to a relative position. The subject judged the prominence of each vowel for each vowel complex. McAdams (21) found that FM increased the prominence of the target vowel. The amount of increase in prominence was greatest when the target vowel was in the highest position (B^b3).

Forrest and Green (10) found a minimum in the Temporal Modulation Transfer Function (TMTF) at a modulation frequency of 10 Hz. McAdams (21) used frequency modulation of about 6 Hz (there was also statistical jitter superimposed on the modulation to mimic voice output). Yost and his co-workers (37, 39) found that there was no difference in detection of SAM for 2, 5, 10 or 20 Hz. There is agreement, therefore, among studies that perturbations in this low frequency region, superimposed upon carriers of higher frequencies, can produce salient acoustic objects.

Temporal Relations Between Signal and Masker

In gap-experiments the task of the subject is to detect the presence of a temporal gap in the stimulus. The gap is an alteration in signal amplitude, a kind of one-time modulation. Carlyon (6) reported that a 250-Hz signal required a larger temporal gap for detection than a 2-kHz signal. His interpretation was that the displacement of the basilar membrane continued for the 250-Hz signal due to ringing while the displacements for the 2-kHz signal died away quickly. The effect of temporal gaps has also been studied in the context of masking experiments. If a gap is produced in a continuing masking noise, the detectability of a signal immediately after the gap is poorer than just prior to the gap, i.e., after the noise has been continuous. The increase in masking, i.e., the decrease in detectability, associated with placing the signal in temporal proximity to masker onset is called <u>overshoot</u>. After some 300 to 500 ms following masker onset, the masking effect of the noise is equivalent to the masking produced by continuous noise, i.e., overshoot diminishes.

McFadden (26) arranged to interrupt either a center band, i.e., a noise band surrounding the signal frequency, or flanking bands, above and below the signal frequency, in order to determine whether a frequency component was associated with the overshoot. With all three bands interrupted, McFadden (26) obtained the classical results: about 10-dB overshoot. When the center band was interrupted while leaving the flanking bands continuous, the subjects showed no overshoot. However, interruption of either flanking band restored the phenomenon. More overshoot was produced by interrupting the upper flanking band than the lower, but both contributed.

Apparently, the time constant of the filter, inferred by Carlyon (6) from his results at 250 Hz and 2 kHz, depends on events occurring at neighboring locations. McFadden (26) measured masking at 4 ms and 300 ms after masker onset. Carlyon's data at 250 Hz showed that a gap of 18 ms was required for detection. The overlap of time values suggests that the time constant of the auditory filter may depend upon events at locations above and below the signal frequency.

Localization of Sound

Interest in the dependence of auditory discrimination upon energy in broad spectral bands has also included work on localization and lateralization of sound. These studies have led to the synthesis of auditory space. Batteau (1) pointed out that the pinna altered the power spectrum of the sound at the entrance to the auditory canal. Blauert (3) and his coworkers measured spectra at the ear canal entrance and Mehrgardt and Mellert (27) made clear that the transfer function from the free sound field to the ear-canal entrance contains the spectral information about direction. Wightman, Kistler and Perkins (36) determined the transfer functions for 144 source positions in an anechoic chamber which included elevations and azimuths. These functions were then used to modify the spectrum of a signal delivered through earphones to each ear to produce spectra corresponding to a specific location in space. Thus, the input signal originating from a given location in space was synthesized for the subject wearing earphones.

Blauert (3) makes the point that the addition of the transfer functions for earphones, ear canals, and the space within which the basic acoustic measurements are made all represent linear phenomena. The transfer functions can be added and their sum provides a filter through which a complex signal might be passed in order to produce auditory experience that duplicates the original. Thus, provided that measurements are made over a representative frequency range it should be possible to synthesize one's favorite music in concert halls of choice. The acoustic pressure measurements must, of course, be made in the specific concert hall at a specific location (seat) in order to capture an acoustic representation of the hall's important spatial features.

With the recognition that broad bandwidths contain cues to source locations, workers found that the traditional cues for azimuth angle, interaural time and intensity differences, had to be considered not just for sinusoids but also for broad spectra. And, for median plane localization, alterations of acoustic energy in selected frequency regions of a broad band noise were found to correlate with subject-assigned elevations (16). The spectral alterations due to source location are produced by resonances and cancellations within the pinna and by reflections from the head and shoulders, depending on the elevation and azimuth of the source (20). Indeed, Wightman, Kistler & Perkins (36) stripped phase information from their spectral representations, leaving only intensity X frequency as the basis for synthesizing location with their procedures. They report correlations exceeding 0.95 between subjects' designations of real and synthesized sources.

It seems unlikely, however, that the auditory system would fail to make use of a cue as prominent as interaural time differences. Since time differences are least ambiguous at low frequencies and intensity differences are most effective at high frequencies, one might presume that the auditory system can parse localization cues across a wide spectrum of acoustic energy. The perception of location and the selection of auditory objects from acoustic backgrounds must be the product of the system's spectral and temporal analysis.

The laboratory findings, reviewed above, that frequencies outside the CB can alter the detection of signals suggests that the auditory system extracts relative differences in acoustic energy among frequency components of the spectrum. If a temporal order, e.g., amplitude modulation, is imposed upon spectral components, the commonality among components is recognized by the auditory system as figure against the acoustic background. Either ear will suffice for the detection of auditory objects and thus, spectra and temporal orders can be processed monaurally. When the second ear is available, the differences between spectra are extracted and used to localize sounds. By processing interaural differences over a wide frequency range, the auditory system reduces ambiguities. For example, interaural time differences are represented in both the front and rear auditory fields; i.e., one interaural time difference may refer to either of two locations. However, the pinna placement helps clarify source location by filtering high frequencies differently, depending on source location. The combination of interaural time differences plus intensity differences helps differentiate front from rear sources.

The folds and creases of the pinna create the intensity variations as a function of source location. For median plane locations, i.e., elevations, alteration in the spectra occur due to the reflections and phase cancellations occurring within the pinna. For example, Hebrank and Wright (16) show that a frontal elevation cue, consisting of a one-octave "notch" or decrease in power, with a lower cut-off frequency that increases with elevation, is related to their subjects' designation Front. The lower frequency of the notch increases from 4 kHz to 8 kHz with elevation and, along with that, there is increased energy above 13 kHz. The notch is created by cancellation due to interference between incident sound and sound reflected from the posterior wall of the pinna. Their designation, Above, was associated with a 1/4 octave peak between 7 and 9 kHz. The reflections from shoulders and torso also contribute to the resultant sound that arrives at the auditory canal.

Synthesis of Auditory Space

There are two levels of interest in the synthesis of auditory space. One is for demonstration and entertainment purposes and the second is for the use of synthesized auditory space as a framework within which information useful for a particular task can be presented. The demonstration level has been attained already; the utility level is still to be achieved. To synthesize auditory space, the power spectrum at each ear for a specific location must be represented in the spectrum of the signal to be localized. The spectrum at one ear produced by a sound from a given location can be expressed in the time domain by a broad-band pulse. The pulse can be convolved with the spectrum representing the acoustic energy in the signal, and their product will be the pressure at that ear synthesized for the specific source location. The same operation may be carried out for the other ear; the two spatially-filtered signals are then presented to both ears simultaneously to produce one localized percept: the signal at the selected location. Since the head position can vary even though a sound source may remain stationary, the broad-band pulses, time-domain representations of different spatially-related spectra, must be selected as the head turns, and convolved with the signal, just as would be necessary to synthesize a moving source. Because of the relation between head position and spatially-representative spectra, there must be some provision for tracking head position in order to select the appropriate pair of spatially-related pulses. The selection and multiplication of the spatially-related broad-band pulses with the incoming sound must be updated quite rapidly to carry out the synthesis in real time, i.e., as the head turns. The bandwidth of the incoming signal also imposes a speed requirement. Many of the demonstrations play music through the system and a magnetic head tracker provides information by which appropriate spatially-related pulses (filters) are selected as the head turns, to keep the sound in the same external position. The processing demanded by the requirement of real time can only be achieved by very high speed computers or, better, by special purpose computers built with high speed chips to carry out operations at megahertz rates. Indeed, the limitations may lie in the slow response of the magnetic head tracker that is now used.

To present useful information within auditory space, there must be some identification of sound with data. For our present application, flight parameters of the aircraft will be associated with perceptual dimensions of the sound. Spatial locations of sound objects may be particularly relevant since the pilot must maintain spatial orientation. The auditory objects might be defferentiated by amplitude modulating some frequencies, by increasing the intensity of some components, etc., following the lead of studies reviewed above.

The relations between stimulus parameters and salience or detectability is the object of study in much of the contemporary research in psychoacoustics. A unifying feature among the papers reviewed is the importance of complex signals as a basis for establishing the subtle discriminations of which the auditory system is capable. When many frequency components are simultaneously present, a wide variety of auditory sensations can be produced by varying component intensities, component frequencies, by modulating component frequencies, or by other means. One study examined the resolution in synthesized auditory space for sounds with the same or different timbres. Divenyi and Oliver (8) used sinusoids, frequency modulated, amplitude modulated and also noise stimuli.

Stimuli were presented simultaneously from two speakers. Subjects were asked to differentiate location (when timbre was the same) or timbre (when location was the same). Divenyi and Oliver reported that the smallest separation between the two speakers that their subjects could discriminate in the horizontal plane was 18 deg; for most sounds presented simultaneously, the subjects required 60 degrees separation. They suggested that when there is spectral overlap, assignment of spatial separation is difficult. Their results suggest that care is required in designing a level of salience into synthesized sounds equal to the spatial resolution of the auditory system.

Presentations of synthesized locations in auditory space could be accompanied by a corresponding synthesized visual field to duplicate the perception of sounds in real three-dimensional space. One would expect that presentation of congruent visual and auditory space would improve the verisimilitude in simulators, etc.

LITERATURE REVIEWED

Most journal articles for this review were taken from the Journal of the Acoustical Society of America. The emphasis is on recent research and the target was to abstract all papers from 1985 forward. There are earlier papers, considered germinal that are also included, as well as books. Even with these specified targets, some papers were probably overlooked, but I estimate that 95% of the literature on these topics was examined. The reference list follows in Appendix A. The bibliography in Appendix B is a complete list of articles reviewed and includes papers referred to in Appendix A. The descriptors following the journal citations represent the topics reviewed above as follows:

CMRELEASE: COMODULATION MASKING RELEASE
LOCALIZ : LOCALIZATION
LATERALIZ: LATERALIZATION
BINAURAL
SPECTRAL
TEMPORAL
FILTER
MODULATION
CORRELATION
UNCERTAINTY

APPENDIX A

REFERENCES

REFERENCES

1. Batteau, D.W. The role of the pinna in sound localization. Proc Roy Soc (Lond. B) 158: 158-180 (1967).
binaural/localiz

2. Bekesy, G. V. Experiments in Hearing. New York: McGraw-Hill Book Company, Inc., 1960.

3. Blauert, J. Spatial Hearing. Cambridge, MA: MIT Press, 1983.

4. Boring, E. G. A history of experimental psychology, 2nd ed.New York: Appleton-Century Crofts, Inc., 1950.

5. Boring, E. G. Sensation and perception in the history of experimental psychology. New York: Appleton-Century-Crofts, Inc., 1942.

6. Carlyon, R. P. The development and decline of forward masking. Hearing Res. 32: 65-80 (1988).
temporal

7. Cohen, M. F. and E. D. Schubert. The effect of cross-spectrum correlation on the detectability of a noise band. J Acous Soc Amer March 81: 721-723 (1987).
cmrelease

8. Divenyi, P.L. and S. K. Oliver. Resolution of steady-state sounds in simulated auditory space. J Acous Soc Amer 85: 2042-2052 (1989).
localiz/binaural/spectral

9. Farrar, C. L., et al. Spectral-shape discrimination. I. Results from normal-hearing listeners for stationary broadband noises. J Acous Soc Amer 81: 1085-1092 (1987).
filter/spectral

10. Forrest, T. G. and D. M. Green. Detection of partially filled gaps in noise and the temporal modulation transfer function. J Acous Soc Amer 82: 1933-1943 (1987).
temporal/filter

11. Green, D. M. Profile Analysis: Auditory Intensity Discrimination. New York: Oxford University Press, Inc., 1988.

12. Hafter, E. R. and L. A. Jeffress. Two-image lateralization of tones and clicks. J Acous Soc Amer 47: 1041-1047 (1968).
lateraliz/binaural

13. Hall, J. W., III, J. A. Cokely, and J. H. Grose. Combined monaural and binaural masking release. J Acous Soc Amer 83: 1839-1845 (1988).
cmrelease/binaural

14. Hall, J. W., III and J. H. Grose. Comodulation masking release: Evidence for multiple cues. J Acous Soc Amer 84: 1669-1675 (1988).
cmrelease

15. Hall, J. W., M. P. Haggard and M. A. Fernandes. Detection in noise by spectro-temporal pattern analysis. J Acous Soc Amer 76: 50-56 (1984).
cmrelease/temporal

16. Hebrank. J. and D. Wright. Spectral cues used in the localization of sound sources on the median plane. J Acous Soc Amer 56: 1829-1834 (1974).
localiz/spectral

17. Hirsh, I.J. The influence of interaural phase on interaural summation and inhibition. J Acous Soc Amer 20: 536-544 (1948).
binaural

18. Klump, R. G. and H. R. Eady. Some measurements of interaural time difference thresholds. J Acous Soc Amer 28: 859-860 (1956).
binaural

19. Koehnke, J. and M. F. Cohen. Masking effects in binaural detection and interaural time discrimination. J Acous Soc Amer 81: 724-729 (1987).
binaural/cmrelease

20. Kuhn, G. F. Physical acoustics and measurements pertaining to directional hearing, ch. 1, pp. 3-25 In W. A. Yost and G. Gourevitch (eds.). Directional Hearing. New York: Springer-Verlag, 1987.
binaural/spectral

21. McAdams, S. Segregation of concurrent sounds.I: Effects of frequency modulation coherence. J Acous Soc Amer 86: 2148-2159 (1989).
modulation/temporal/spectral

22. McFadden, D. Comodulation masking release:Effects of varying the level, duration, and time delay of the cue band. J Acous Soc Amer 80: 1658-1667 (1986).
cmrelease

23. McFadden, D. Comodulation detection differences using noise-band signals. J Acous Soc Amer 81: 1519-1527 (1987).
cmrelease/modulation

24. McFadden, D., L. A. Jeffress, and H. L. Ermey. Differences of interaural phase and level in detection and lateralization: 250 Hz. J Acous Soc Amer 50(Part 2): 1384-1493 (1971).
lateraliz/binaural

25. McFadden, D., L. A. Jeffress, and J. R. Lakey. Differences of interaural phase and level in detection and lateralization: 1000 and 2000 Hz. J Acous Soc Amer 52(Part 2): 1197-1206 (1972).
binaural/lateraliz

26. McFadden, D. Spectral differences in the ability of temporal gaps to reset the mechanisms underlying overshoot. J Acous Soc Amer 85: 254-261 (1989).
filter/temporal

27. Mehrgardt, S. and V. Mellert. Transformation characteristics of the external human ear. J Acous Soc Amer 61: 1567-1576 (1977).
binaural/spectral

28. Patterson, R. D. A pulse ribbon model of monaural phase perception. J Acous Soc Amer 82: 1560-1586 (1987).
spectral/temporal

29. Patterson, R. D. Auditory filter shapes derived with noise stimuli. J Acous Soc Amer 59:640-654 (1976).
spectral

30. Plenge, G. On the difference between localization and lateralization. J Acous Soc Amer 56: 944-951 (1974).
localiz/spectral

31. Richards, V. M.. Monaural envelope correlation perception. J Acous Soc Amer 82: 1621-1630 (1987).
cmrelease/correlation

32. Schooneveldt, G. P. and B. C. J. Moore. Comodulation masking release (CMR): Effects of signal frequency, flanking-band frequency, masker bandwidth, flanking-band level, and monotic versus dichotic presentation of the flanking band. J Acous Soc Amer 82: 1944-1956 (1987).
cmrelease/binaural

33. Stevens, S. S. and E. B. Newman. The localization of actual sources of sound. Amer J Psychol 48: 297-306 (1936).
localization

34. Watson, C. S., et al. Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals. J Acous Soc Amer 57: 1175-1185 (1975).
temporal/spectral

35. Watson, C. S., W. J. Kelly, and H. W. Wroton. Factors in discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. J Acous Soc Amer 60: 1176-1186 (1976).
temporal/spectral/uncertainty

36. Wightman, F. L., D. J. Kistler, and M. E. Perkins. A new approach to the study of human sound localization, ch. 2, pp 26-48. In W. A. Yost and G. Gourevitch (eds.). Directional Hearing. New York: Verlag-Verlag, 1987.
binaural/localiz

37. Yost, W.A., & Sheft, S. Across-critical band processing of amplitude-modulated tones. J Acous Soc Amer 85: 848-857 (1989).

38. Yost, W. A. and E. R. Hafter. Lateralization, ch 3, pp. 49-84. In W. A.Yost and G. Gourevitch (eds.). Directional Hearing. New York: Springer-Verlag, 1987.
binaural/lateraliz

39. Yost, W. A., S. Sheft, and J. Opie. Modulation interference in detection and discrimination of amplitude modulation. J Acous Soc Amer 86: 2136-2147 (1989).
modulation

APPENDIX B

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Algom, D., R. Adam, and L. Cohen-Raz. Binaural summation and lateralization of transients: A combined analysis. J Acoust Soc Amer 83:1302-1316 (1988).
   binaural/lateraliz.

2. Amenta, C.A., III, et al. Some physical and psychological effects produced by selective delays of the envelope of narrow bands of noise. Hearing Res 29:147-161 (1987).
   binaural/lateraliz

3. Bacon, S. P. and B. C. J. Moore Transient masking and the temporal course of simultaneous tone-on-tone masking. J Acous Soc Amer 81:1073-1077 (1987).
   temporal

4. Bernstein. L.R. and D. M. Green, Detection of simple and complex changes of spectral shape. J Acous Soc Amer 82:1587-1592 (1987).
   profile

5. Bernstein, L. R. and D. M. Green The profile analysis bandwidth. J Acous Soc Amer 81:1888-1895 (1987).
   profile/filter/binaural

6. Bregman, A.S., et al. Spectral integration based on common amplitude modulation. Percept and Psychophys 37: 483-493 (1985).
   modulation.

7. Buell, T.N. and E. H. Hafter. Discrimination of interaural differences of time in the envelopes of high-frequency signals: Integration times. J Acoust Soc Amer 84: 2063-2066 (1988).
   lateraliz/temporal

8. Carlyon, R.P. A release from masking by continuous, random, notched noise. J Acous Soc Amer 81: 418-426 (1987).
   filter/adaptation

9. Cohen, M.F. and E. D. Schubert. Influence of place synchrony on detection of a sinusoid. J Acous Soc Amer 81: 452-458 (1987).
   cmrelease/binaural

10. Deng, L. and C. D. Geisler. Responses of auditory-nerve fibers to multiple-tone complexes. J Acous Soc Amer 82: 1989-2000 (1982).
    neural/profile/modulation

11. Dooley, G.J. and B. C. J. Moore. Detection of linear frequency glides as a function of frequency and duration. J Acous Soc Amer 84: 2045-2057 (1984). temporal.

12. Dooley, G. J. and B. C. J. Moore. Duration discrimination of steady state and gliding tones: A new method for estimating sensitivity to rate of change. J Acous Soc Amer 84: 1332-1337 (1988).
temporal

13. Durlach, N. I. Equalization and Cancellation theory of binaural masking level differences. J Acous Soc Amer 35: 1206-1218 (1963).
binaural/localiz

14. Fantini, D. A. and D. S. Emmerich. Edge effects on frequency discrimination of tones presented in low- and high-pass noise backgrounds. J Acous Soc Amer 82:1593-1597 (1987).
filter

15. Formby, C. and K. Muir. Modulation and gap detection for broadband and filtered noise signals. J Acoust Soc Amer 84: 545-550 (1988).
temporal/filter

16. Geisler, C.D. and T. Gamble. Responses of "high-spontaneous" auditory-nerve fibers to consonant-vowel syllables in noise. J Acous Soc Amer 85: 1639-1652 (1989).
neural/spectral

17. Grantham, D. W. and L. E. Luethke. Detectability of tonal signals with changing interaural phase difference in noise. J Acous Soc Amer 83: 1117-1123 (1988).
lateraliz/binaural/temporal

18. Green, D.M,, A. O. Zekiye, and T. Forrest. Frequency effects in profile analysis and detecting complex spectral changes. J Acous Soc Amer 81: 692-699 (1987).
profile

19. Green, D.M. and T. G. Forrest. Temporal gaps in noise and sinusoids. J Acoust Soc Amer 86: 961-970 (1989).
temporal

20. Grose, J.H. and J. W. Hall, III. The effect of signal-frequency uncertainty on comodulation masking release. J Acous Soc Amer 87: 1272-1277 (1990).
cmrelease

21. Grose, J. H., D. A. Eddins, and J.W. Hall, III. Gap detection as a function of stimulus bandwidth with fixed high-frequency cutoff in normal-hearing and hearing-impaired listeners. J Acous Soc Amer 86: 1747-1755 (1989).
temporal

22. Hafter, E. R., et al. The combination of interaural time and intensity in the lateralization of high-frequency complex signals. J Acous Soc Amer 87: 1702-1708 (1990).
lateraliz/binaural

23. Hall, J.W., III, J. H. Grose, and M. P. Haggard. Comodulation masking release for multicomponent signals. J Acous Soc Amer 83: 677-686 (1988).
cmrelease/correlation

24. Hall, J. W., III, et al. Spectro-temporal analysis in normal-hearing and cochlear-impaired listeners. J Acous Soc Amer 84: 1325-1331 (1988).
cmrelease

25. Hartmann, W. M. and B. Rakerd. On the minimum audible angle--A decision theory approach. J Acoust Soc Amer 85: 2031-2041 (1989).
localiz

26. Hartmann, W. M. and B. Rakerd. Localization of sound in rooms IV: The Franssen effect. J Acous Soc Amer 86: 1366-1373 (1989).
localiz/temporal

27. Hebrank, J. and D. Wright. Are two ears necessary for localization of sound sources on the median plane? J Acous Soc Amer 56: 935-938 (1974).
localiz/spectral

28. Humanski, R. A. and R. A. Butler. The contribution of the near and far ear toward localization of sound in the sagittal plane. J Acous Soc Amer 83: 2300-2310 (1988).
localiz/spectral

29. Kidd, G., Jr., C. R. Mason, and T. E. Hanna. Evidence for sensory-trace comparisons in spectral shape discrimination. J Acous Soc Amer 84: 144-149 (1988).
profile/spectral

30. Kohlrausch, A. Auditory filter shape derived from binaural masking experiments. J Acoust Soc Amer 84: 573-583 (1988).
binaural/filter

31. Kollmeier, B. and R. H. Gilkey. Binaural forward and backward masking: Evidence for sluggishness in binaural detection. J Acous Soc Amer 87: 1709-1719 (1990).
binaural/temporal/filter

32. Lutfi, R. A. Informational processing of complex sound. I: Intensity discrimination. J Acous Soc Amer 86: 934-944 (1989).
profile/uncertainty

33. Lutfi, R.A. Informational processing of complex sound. II. Cross-dimensional analysis. J Acous Soc Amer 87: 2141-2148 (1990).
profile

34. Makous, J. C. and J. C. Middlebrooks. Two-dimensional sound localization by human listeners. J Acous Soc Amer 87: 2188-2200 (1990).
localiz/binaural/spectral

35. May, B., D. B. Moody, and W. C. Stebbins. Categorical perception of conspecific communication sounds by Japanese Macaques, Macaca fuscata. J Acous Soc Amer 85: 837-847 (1989).
spectral

36. McFadden, D. and B. A. Wright. Comodulation masking release in a forward-masking paradigm. J Acous Soc Amer 82: 1615-1620 (1987).
cmrelease/temporal

37. Middlebrooks, J. C., J. Makous, and D. M. Green. Directional sensitivity of sound-pressure levels in the human ear canal. J Acous Soc Amer 86: 89-108 (1989).
localiz/binaural/spectral

38. Middlebrooks, J.C. and D. M. Green. Directional dependence of interaural envelope delays. J Acous Soc Amer 87: 2149-2162 (1990).
localiz/binaural/spectral

39. Moore, B.C.J., B. R. Glasberg, and G. P. Schooneveldt. Across-channel masking and comodulation masking release. J Acous Soc Amer 87: 1683-1694 (1990).
cmrelease/temporal

40. Moore, B. C. J., S. R. Oldfield, and G. J. Dooley. Detection and discrimination of spectral peaks and notches at 1 and 8 kHz. J Acous Soc Amer 85: 820-836 (1989).
spectral

41. Moore, B. C. J. and B. R. Glasberg. Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes. J Acous Soc Amer 82: 69-79 (1987).
cmrelease/temporal

42. Moore, B. C. J. and B. R. Glasberg. Frequency discrimination of complex tones with overlapping and non-overlapping harmonics. J Acous Soc Amer 87: 2163-2177 (1990).
spectral

43. Moore, B. C. J. and B. R. Glasberg. Mechanisms underlying the frequency discrimination of pulsed tones and the detection of frequency modulation. J Acoust Soc Amer 86: 1722-1732 (1989).
spectral/filter

44. Moore, B. C. J. and D.S. Emmerich. Monaural envelope correlation perception, revisited: Effects of bandwidth, frequency separation, duration, and relative level of the noise bands. J Acous Soc Amer 87: 2628-2633 (1990).
cmrelease

45. Moore, B. C. J., et al. The shape of the ear's temporal window. J Acous Soc Amer 83: 1102-1116 (1988).
filter/temporal

46. Moore, B. C. J. and B. R. Glasberg. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. J Acous Soc Amer 74: 750-753 (1983).
filter

47. Moore, B. C. J., et al. The temporal course of masking and the auditory filter shape. J Acous Soc Amer 81: 1873-1880 (1987).
temporal/filter

48. Moore, Brian C. J. An introduction to the psychology of hearing. New York: Academic Press, 1982.

49. Neff, D. L. and B. P. Callaghan. Effective properties of multicomponent simultaneous maskers under conditions of uncertainty. J Acous Soc Amer 83: 1833-1838 (1988).
profile/uncertainty

50. Noble, W. Auditory localization in the vertical plane: Accuracy and constraint on bodily movement. J Acous Soc Amer 82: 1631-1636 (1987).
localis/binaural

51. Ozimek, E. and A. Sek. Perception of amplitude and frequency modulated signals (mixed modulation). J Acous Soc Amer 82: 1598-1603 (1987).
modulation

52. Perrot, D. R., et al. Minimum audible angle thresholds obtained under conditions in which the precedence effect is assumed to operate. J Acous Soc Amer 85: 282-288 (1989).
localiz/temporal

53. Perrot, D. R. and J. Tucker. Minimum audible movement angle as a function of signal frequency and the velocity of the source. J Acous Soc Amer 83: 1522-1527 (1988).
localiz

54. Perrot, D. R. and S. Kourosh. Minimum audible angle thresholds for sources varying in both elevation and azimuth. J Acous Soc Amer 87: 1728-1731 (1990).
localiz

55. Perrott, D. R. and S. Pacheco. Minimum audible angle thresholds for broadband noise as a function of the delay between the onset of the lead and lag signals. J Acous Soc Amer 85:2669-2672 (1989).
localiz/temporal

56. Perrott, D. R., H. Abarsoom, and J. Tucker. Changes in head position as a measure of auditory localization performance: Auditory psychomotor coordination under monaural and binaural listening conditions. J Acous Soc Amer 82: 1637-1645 (1987).
localiz/binaural

57. Pickles, J. O. An Introduction to the Physiology of Hearing. New York: Academic Press, 1982.

58. Plack, C. J. and B. C. J. Moore. Temporal window shape as a function of frequency and level. J Acous Soc Amer 87: 2178-2187 (1990).
temporal/filter

59. Preece, J. P. and R. H. Wilson. Detection, loudness, and discrimination of five-component tonal complexes differing in crest factor. J Acous Soc Amer 84: 166-171 (1988).
profile

60. Puel, J-L, and G. Rebillard. Effect of contralateral sound stimulation on the distortion product 2F1-F2:Evidence that the medial efferent system is involved. J Acous Soc Amer 87: 1630-1635 (1990).
neural

61. Raney, J. J., et al. Signal uncertainty and psychometric functions in profile analysis. J Acous Soc Amer 86: 954-960 (1989).
profile

62. Saberi, K. and D. R. Perrot. Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate. J Acous Soc Amer 87: 1732-1737 (1990).
localiz

63. Schneider, B. A. and P. M. Zurek. Lateralization of coherent and incoherent targets added to a diotic noise background. J Acous Soc Amer 83: 1756-1763 (1988).
lateraliz/binaural

64. Schneider, B. A., D. Bull, and S. E. Trehub. Binaural unmasking in infants. J Acous Soc Amer 83: 1124-1132 (1988).
binaural/localiz

65. Schooneveldt, G. P. and B. C. J. Moore. Failure to obtain comodulation masking release with frequency-modulated maskers. J Acous Soc Amer 83: 2290-2292 (1988).
cmrelease

66. Shailer, M. J. and B. C. J. Moore. Gap detection and the auditory filter: Phase effects using sinusoidal stimuli. J Acous Soc Amer 81: 1110-1117 (1987).
filter/temporal

67. Shamma, S. A., N. Shen, and P. Gopalaswamy. Stereausis: Binaural processing without neural delays. J Acous Soc Amer 86: 989-1006 (1989).
neural/binaural/temporal

68. Shaw, E. A. G. Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. J Acous Soc Amer December 56: 1848-1861 (1974).
localiz/spectral

69. Sidwell, A. Spectral and level effects in noise-on-tone suppression. J Acous Soc Amer 81: 1078-1084 (1987).
temporal/filter

70. Siegal, R. A. and H. S. Colburn. Binaural processing of noisy stimuli: Internal/external noise ratios for diotic and dichotic stimuli. J Acous Soc Amer 86: 2122-2128 (1989).
binaural

71. Small, A. M. Lateralization of dichotic noise bursts: Effect of onset and offset disparity. J Acous Soc Amer 82: 1957-1966 (1987).
lateraliz/temporal

72. Sorkin, R. D. Perception of temporal patterns defined by tonal sequences. J Acous Soc Amer 87: 1695-1701 (1990).
temporal/uncertainty

73. Sorkin, R. D. Temporal factors in the discrimination of tonal sequences. J Acous Soc Amer 82: 1218-1226 (1987).
temporal/correlation/uncertainty

74. Staffel, J. G., et al. NoSo and NoSii detection as a function of masker bandwidth in normal-hearing and cochlear-impaired listeners. J Acous Soc Amer 87: 1720-1727 (1990).
binaural

75. Stern, R. M., A. S. Zeiberg, and C. Trahiotis. Lateralization of complex binaural stimuli: A weighted-image model. J Acous Soc Amer 84: 156-165 (1988).
lateraliz/spectral

76. Strickland, E. A., et al. Within- versus cross-channel mechanisms in detection of envelope phase disparity. J Acous Soc Amer 86: 2161-2166 (1989).
filter/binaural/spectral

77. Summerfield, Q., A. Sidwell, and T. Nelson. Auditory enhancement of changes in spectral amplitude. J Acous Soc Amer 81: 700-707 (1987).
adaptation/spectral

78. Tomlinson, R. W. W. and D. W. F. Schwarz. Perception of the missing fundamental in nonhuman primates. J Acous Soc Amer 84: 560-565 (1988).
spectral

79. Trahoitis, C., and R. M. Stern. Lateralization of bands of noise: Effects of bandwidth and differences of interaural time and phase. J Acous Soc Amer 86: 1285-1293 (1989).
lateraliz

80. Webster, F. A., The influence of interaural phase on masked thresholds: .I The role of interaural time deviation. J Acous Soc Amer 23: 452-462 (1951).
binaural

81. Wightman, F., and D. J. Kistler. Headphone simulation of free-field listening. I:Stimulus Synthesis. J Acous Soc Amer 85: 858-867 (1989).
localiz/binaural/spectral

82. Wightman, F.L. and D. J. Kistler. Headphone simulation of free-field listening. II:Psychophysical validation. J Acous Soc Amer 85: 868-878 (1989).
localiz/binaural/spectral

83. Yost, W. A. and G. Gourevitch. Directional Hearing. New York: Springer-Verlag Inc., 1987.

84. Yost, W. A. and R. H. Dye, Jr. Discrimination of interaural differences of level as a function of frequency. J Acous Soc Amer 83: 1846-1851 (1988).
lateraliz/binaural

85. Yost, W. A. and M. J. Moore. Temporal changes in a complex spectral profile. J Acous Soc Amer 81: 1896-1905 (1987).
profile/modulation

86. Zatorre, J. Pitch perception of complex tones and human temporal-lobe function. J Acous Soc Amer 84: 566-572 (1988).
neural

87. Zurek, P. M. and N. I. Durlach. Masker-bandwidth dependence in homophasic and antiphasic tone detection. J Acous Soc Amer 81: 459-464 (1987).
binaural/filter

* U S. GOVERNMENT PRINTING OFFICE 1991--561-052/4000